



[illegible]

submitted to the Protein Sequence Database, May 1996

A:Reference number: S64071

A:Accession: S64243

A:Molecule type: DNA

A:Residues: 1-288 <RID>

A:CROSS-references: EMBL:J272743; NID:g1322868; PID:CAA96937.1; PID:g1322869; MIPS:YQ

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SGD:NIF3

A:CROSS-references: SGD:S0003189; MIPS:YGL221c

A:Map position: 7L

C:Superfamily: conserved hypothetical protein YGL144c

Alignment Scores:

Pred. No.:	2_31e-24	Length:	288
Score:	384.50	Matches:	104
Percent Similarity:	44.78%	Conservative:	59
Best Local Similarity:	28.57%	Mismatches:	98
Query Match:	13.79%	Indels:	103
DB:	2	Gaps:	11

US-09-745-506-74 (1-1553) x S64243 (1-288)

```
OY      251 TTGAAGCTCTCCCTTTCTTCCTTGATGACATTGGCATCCCTCGTGTCTGAC---AGT 307
          ||| -|||:: ||::: |||
Db       10 LeuAaprlsLeuValArgSerlleThrLysPheTyProGlnLysTryAlaAspLysSer 29
          ||| ||||| |
OY      308 TGGGCACATGTTGGATTACTGTGGTAGCAACAACCACACACATACGTAAATVACA----- 361
          ||||| ||||| ||||| ||||| ::|
Db       30 TrpApsanThrGlyLeuLeuIleAspYserThrAlaGlnValThrThrAlaAspAla 49
          ||||| ||||| ||||| ||||| :|
OY      362 -----CTCTTCCTGACCAATGACCTGACTGTGAGGAAGTGAGGAGCTG 406
          ::| ||||| ||||| ||||| |||
Db       50 AsnAlatylThrLysValLeuLeuThrValAspLeuThrLysSerValAlaGlnGluAla 69
          ||| ||||| ||||| ||||| |||||
OY      407 CTGCCAAAAGAAGSCAGCATCTATCTCTCCATACATCCGCGCTATCTTCCGACCATGAG 466
          ::: ||||| ||||| ||||| ||||| |||||
Db       70 ValAapAlaAsnCysAsnValIleValAlaTrpHisProMetIlePheProSetrTrpAsn 89
          ||||| ||||| ||||| ||||| |||||
OY      467 CGCATTAACCTGGACACACA--TGGAAGAGCGCGCTGTGATCCGGGCTCTGGAAACAGA 523
          |||||::: ||||| ::|
Db       90 ArgLeuSerProHistiThrAsnProGlnHisGluThrAlaIleLysLeuIleGlnTyGly 109
          |||||::: ||||| ::|
OY      524 GTGGGTATCTACTCTCCATPACAGCCCTATGATGTGGCCCCAGGGCGTCAACAACGTGG 583
          ::::: ||||| ||||| ||||| ||||| |||||
Db       110 IleSerValTYCysProHistiThrAlaValAspAlaAlaArgGlyGlyValAsnAspTrp 129
          ||||| ||||| ||||| ||||| |||||
OY      584 TTGGCTAAAGGGCTTGGAGCTGTACTCCAGGCCCATACATCCTCCMAAGCTCCCAAC 643
          ||| ::||| |||||
Db       130 LeuValAlaArgGlyLeuAsn----- 135
          |||
OY      644 TAACCTACAGAGGGAACACCAGGTAGATAATTCAAGCTTAACTACACCCAAGACCTGGAC 703
          |||
Db       135 ----- 135
          |||
OY      704 AAAGTCATGTCTGAGGTGAAGGATGTGACGGSTGTTCTGTCACHT---TCTTTTCTGCT 760
          ||| |||||
Db       136 -----asnGlyIuaSnValAlaIalysSerTyAlaIeu 146
          ||| |||||
OY      761 AGGACTGTGAATGAGAGAACAAACAGGATTATCTGAATTGTACTCAGAAAGCTTTGATG 820
          ||| ::| |||
Db       147 GluThrValSerGlyGlu----- 152
          |||
OY      821 CAGGTGTGATGATTTTCTTCCCGGACAACAACACTTATCAGAAAGCGGAATTTCTGICA 880
          |||::: |||
Db       153 -----ThlAspAspLeu--- 156
          |||
OY      881 CTGAGGAAGCCTTGTGCTTCTACATAGTGAATGGAGCGATTATGCACACTGGATTAATCT 940
          ||| |||||
Db       157 -----IleelTyTrgIaIgrPheValGluPhaSnIuIyasp 168
          |||
OY      941 GTCTCCCTGGCAACCATGATTGATGCAATTAATAAAGACACCTAAACATATCTCATATTTGCG 1000
          ::||| ||||| ||||| |||||
```







Db 261 LysGlnValLysLysAlaPheAspValProThrAlaArgVal-----ValGly--- 276  
QY 1022 ACCTTAAGTCTCAAGTCAAAAGTGGCCCTGTGTGCTGTGCTGGAGCAGCGTTCTG 1081  
Db 277 SerLeuGlnThrGlnIleArgLysValAlaValLeuGlnLysAspGlyAsnLysTyrMet 296  
QY 1082 CAGGGTGT-----GAGCGTACCTTTACTCTACAGAGTGAAGATGCCCATCATAT 1132  
Db 297 AlaHisAlaLeuArgLysGlyAlaAspValIleValThrGlyAspValTyrHisVal 316  
QY 1133 ACTTGGATGCTGCTCCCAAGAAATAGTCATCCCTGTGGAACAGACAGACCTGAA 1192  
Db 317 ProHisAspAlaLeuMetAspGlyLeuAsnIleValAspProGlnHis---AsnValGlu 335  
QY 1193 CGAGCGTTTCTTCTGACCTTCGAGATATCTGATTCCTGATTCCTGAGAAATAG----- 1246  
Db 336 LysIleMetLysGlnGlyValLysGlnLysLeuGlnLysLeuLysAspLysLysTyr 355  
QY 1247 ---ATTAATATATCTATCTACAGACTGACAGGAGCCTTTACAGTGTGTA 1294  
Db 356 AspThrGlnValValAlaSerSerValHisThrAspProPheThrPheIle 372

RESULT 7  
AH1618  
conserved hypothetical protein lin1489 [imported] - *Listeria innocua* (strain C1p11262)  
C:Species: *Listeria innocua*  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AH1618  
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund, A.; Title: Comparative genomics of *Listeria* species.  
A:Title: Comparative genomics of *Listeria* species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AH1618  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-373 <GLA>  
A:Cross-references: GB:AL592022; PIDN:CAC96720.1; PID:gl6413962; GSPDB:GN00178  
A:Experimental source: strain C1p11262  
C:Genetics:  
A:Gene: lin1489

Alignment Scores:  
Pred. No.: 2,166-13 Length: 373  
Score: 256.00 Matches: 86  
Percent Similarity: 42.93% Conservative: 78  
Best Local Similarity: 22.51% Mismatches: 152  
Query Match: 9.18% Indels: 66  
Gaps: 11

US-09-745-506-74 (1-1553) x AH1618 (1-373)

QY 272 TTGAATGACTTGTGCATCCCTCTGTTGCTGAGACTGGACATGTTGGATTACTGTG 331  
Db 13 MetGlnLysIleAlaProLysLysLeuAlaMetGlnLysAspProIleGlyLeuGlnVal 32  
QY 332 GAACCAAGCCCAACCATACATCTGTAATACACTCTCTGACCAATGACCTGACTAGAA 391  
Db 33 GlysAspLeuSerLysLysValAlaGlyLysValMetPhe---ThrLeuAspValLeuGlnVal 51  
QY 392 GTGATGAGAGAGTGTCTCAAAAGAGAGAGACCATCTTCTCTACATCCGCTATC 451  
Db 52 ValValAspGlnAlaIleGlnLysArgValAspLeuIleIleAlaHisHisProPheLeu 71  
QY 452 TTCGACCCATGAGACGATACCTGGAACACATGGAAGAGAGCGCTGTGATCCGGCT 511  
Db 72 TyrArgProThrGlnHisIleAspThrThrThrLysGlnGlyLysMetIleLysLysLeu 91  
QY 512 CTGGAACACAGAGTGTACTCTCTCTATACAGCTATGATGCTGCGCCCAAGGC 571

Db 92 IleLysHisAspIleThrValPheAlaAlaHisThrAsnLeuAspIleAlaGlnGly 111  
QY 572 GTCAACAATGTTGGTAAAGGCTTGAAGCTGTACTCTCCAGGCCCATCATCTTCC 631  
Db 112 ValAsnAspIleLeuAlaAspLeuLeuHisIleGlnAspThrThrMetIleGlnThr 131  
QY 632 -----AAAGCTCCCACTACCTCACAG----- 655  
Db 132 TyrThrGlnProTyrCysLysIleAlaValIleAlaProGlnAsnGlnLysLeuVal 151  
QY 656 -----GGAACCCAGAGTAGAATTAACGTTAATCACTACACCCAGACCTG 700  
Db 152 ArgLeuAlaLeuValAsnAsnGlyAlaGlyIleGlnIleGlyThrAsnTyrThrGlu----- 169  
QY 701 GACAAGTCAATGTCTGACGTAAGAAATGACAGGTGTCTGTCTCTCTTTTCTGCT 760  
Db 170 -----CysThrPheHisThrThrGlyLysLeuSerPheLysProGlyThrAspAlaAsn 187  
QY 761 AGGACTGTAATGAGAACCAACAGGATTAATCTG----- 796  
Db 188 ProThrIleGlnGlyLysGlnThrLeuThrSerIleProGlnValLysIleGlnAlaIle 207  
QY 797 -----AATGTACTCAGAGGCTTGATG----- 820  
Db 208 PheProGlnTyrLeuThrGlnThrIleThrLysAlaValLysIleAlaHisProTyrGlu 227  
QY 821 ---CAGGTGATGATTTTCTTCCCGAACAACATTTATCAGACAGCAATCTG 877  
Db 228 GlnProAlaIleAspValTyrThrLeuGlnThrGlnThrTyrLysGln----- 243  
QY 878 TCACGTGAGAAACCTTCTCTCTACATCTGCAATGAGGAGCTTATGACACTGTGAGAA 937  
Db 244 -----GlyLeuGlnArgValGlyThrLeuProLys 253  
QY 938 TCTGCTCTCGCAACCATGATGATGCAATTAAGACACCTAAATCTCATATT 997  
Db 254 LysIleSerMetValSerPheIleAspLysLeuLysThrAlaPheAlaIleAspAsnVal 273  
QY 998 CGCTTAGCCCTTGGGTGGGAGAACCTTACAGTCTCAAGTCAAGTCTGCGCCTGT 1057  
Db 274 ArgPhe-----ValGly---AspLeuLysAlaAsnValGlnLysValAlaIleIle 289  
QY 1058 GCTGTTTGGGACGAGCGTTCTG-----CAGGCTGTGAGCGTACCTTACCTC 1108  
Db 290 GlnGlyAspGlyAsnLysPheIleHisGlnAlaLysAlaThrGlyAlaAspValPheIle 309  
QY 1169 CTCTGTGACACAGC-----AACACTGACAGGCTTCTTCTGACCTTCGAGATATG 1222  
Db 330 AspAlaGlnHisAsnIleGlnLysValMetLysGlyTyrLeuLysAsnLysMetLysGlu 349  
QY 1223 CTGGAATCTCACTGGAATTAAGTAAATATATATCACTATGACAGACGAGACCT 1282  
Db 350 GlnAlaLysIleLeuAspTyrGlnAlaGlnPheIleValSerGlnValAsnThrAspPro 369  
QY 1283 CTTCAG 1288  
Db 370 PheGln 371

RESULT 8  
AD1256  
conserved hypothetical protein lmo1452 [imported] - *Listeria monocytogenes* (strain EG C:Species: *Listeria monocytogenes*  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AD1256  
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl, D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapatk, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schlueter, T.; Simes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of *Listeria* species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AD1256  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-373 <G1A>  
 A:Cross-references: GB:NC\_003210; PIDN:CAC99530.1; PID:g16410881; GSPDB:GN00177  
 A:Experimental source: strain EGD-e  
 C:Genetics:  
 A:Gene: Imo1452

## Alignment Scores:

Pred. No.:	4,95e-12	Length:	373
Score:	240.00	Matches:	88
Percent Similarity:	45.97%	Conservative:	83
Best Local Similarity:	23.66%	Mismatches:	155
Query Match:	8.61%	Indels:	46
DB:	2	Gaps:	14

US-09-745-506-74 (1-1553) x AD1256 (1-373)

```

OY 272 TTGAATGACTTTGCAATCCCTCTGTTTCTGAGATTGGACAATGTTGGATTACTGTG 331
DB 13 MetGluLysIleAlaProLysLysLeuAlaMetGluLysProIleGlyLeuGluVal 32
OY 332 GAACCAAGCCACCACTACTGTAATATACACTCTCTCTGACCAATGACCTGACGAGAA 391
DB 33 ---GlyAAspLeuSerArgLysValArgLysIleMetPheThrIleAspValLeuGluGlu 51
OY 392 GTGATGAGAGAGGTGTGCAAAAGAGGACACCTCATTTCTCTCCATCTCCGCTATC 451
DB 52 ValValAspLysIleAlaIleGluLysValAspLeuIleLeuAlaHisPheProPheLeu 71
OY 452 TTCGAGCCATGAGCCCATACCTGAGACACATGAGAGAGCCCTGTGATCCGGCT 511
DB 72 TyrArgProThrGlnHisIleAspThrThrThrGlyGlnGlyMetIleLysLeu 91
OY 512 CTGGAGACAGAGCGGTATCTACTCTCCATACACCATATGATGCTGCGCCACAGGC 571
DB 92 IleLysHisAspIleThrValPheAlaAlaHisThrAsnLeuAspIleAlaGlnGly 111
OY 572 GTCAACAACTGTGTGCTAAAGGCTGTGAGCTGTGATCTCCAGCCCATCATCTTCC 631
DB 112 ValAsnAsnIleLeuAlaAspLeuLeuHisLeuGlnAsnThrMetIleGluGluThr 131
OY 632 AAAGCTCCCACTACCTACAGAGGAAACCCAGTAGAATTCAAGCTTAACATACAC 691
DB 132 TyrSerGluProTyrCysLysIleAlaValTyrValProGlu----- 145
OY 692 CAGACCTGGACAAGATGCTGACGTCG-----AAAGAAATTGACGGCTT 739
DB 146 AsnGluLeuGlnSerValArgLeuAlaLeuValAsnGlnLysGlnIleGlyThr 165
OY 740 TCTGTACTCTTTTCTGTAGACTGT-----AATGAGAAACAACAGGATTAAT 793
DB 166 GluTyrThrGluCysThrPheHisThrThrGlyIleGlySerPheLysProGlyAlaAsn 185
OY 794 CTGAATGTACT-----CAGAGGCTTTGATGCAAGTGTGATTTTCTTCCCG 844
DB 186 AlaAsnProThrIleGlyLysAspAlaLeuThrSerValProGluVal-----Lys 203
OY 845 AACAAACACTTTATCAGAG-----ACGAAATTGTGCA----- 880
DB 204 IleGluAlaIlePheProGlnTyrLeuThrGluThrIleThrLysAlaValLysIleAla 223
OY 881 -----CTGAGAGACCTTGTCTT-----CTACATACT----- 907
DB 224 HisProTyrGluGluProAlaIleAspValTyrThrLeuGluMetGlnThrTyrLysGlu 243
OY 908 GGAATGGAGCGTTATGACACCTGGATGAATCTGTCCCTGGCAACCATGATGATGCA 967
  
```

```

DB 244 GlyLeuGlyArgValGlyMetLeuProLysLysLeuGlyMetValSerPheIleAspLys 263
OY 968 ATAAAAAGACACCTAAACATATCTCATATTCGCTTACCCCTGGGGGGGAGAACCTTA 1027
DB 264 LeuLysThrAlaPheAlaIleAspAsnValArgPhe-----IleGly---AspLeu 279
OY 1028 GAGCTCAAGCAAAAGTCGCGCTGTGCTGCTGCTGGAGACAGCTTCTG----- 1081
DB 280 LysThrThrValHisLysValAlaIleIleGlyLysPheLysAsnLysPheIleHisGln 299
OY 1082 ---CAGGCTGTGAGCTGACCTTACTCAGATGAGATGTCCTCATGATATCTTGG 1138
DB 300 AlaLysSerThrGlyAlaAspValPheIleThrGlyAspValTyrTyrHisThrGlyHis 319
OY 1139 GATGCTCTTCCCAAGGAATTAATGTCATCTCTGTGACACAGC-----AACACTGAA 1192
DB 320 AspLeuLeuAlaIleAsnLeuProThrIleAspAlaGlnHisAsnIleGluLysValMet 339
OY 1193 CGAGCTTTTCTTCTGACCTTCGAGATATGCTGATTCCTCACTGGAATTAATTAAT 1252
DB 340 LysGlyTyrLeuLysAsnLysMetGluGlnAlaLysIleLeuAspTyrGluAlaGlu 359
OY 1253 ATATCTATCAGACTGACAGGACCGCTTTCAG 1288
DB 360 PheIleValSerGluValAsnThrAspProPheGln 371
  
```

RESULT 9

DB1341  
 hypothetical protein Cj0705 [Imported] - Campylobacter jejuni (strain NCTC 11168)  
 C:Species: Campylobacter jejuni  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002  
 C:Accession: DB1341  
 R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chai, C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; Vanyivet, A.; Whitehead, S.; Barr, Nature 403, 665-668, 2000  
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals  
 A:Reference number: A81250; MUID:20150912; PMID:10688204  
 A:Accession: DB1341  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-241 <PAB>  
 A:Cross-references: GB:AL139076; GB:AL11168; MTD:66966128; PIDN:CA872979.1; PID:9696  
 A:Experimental source: serotype O2, strain NCTC 11168  
 C:Genetics:  
 A:Gene: Cj0705

Alignment Scores:

Pred. No.:	7.92e-11	Length:	241
Score:	225.00 <td>Matches:</td> <td>68 </td>	Matches:	68
Percent Similarity:	37.01% <td>Conservative:</td> <td>63 </td>	Conservative:	63
Best Local Similarity:	19.21% <td>Mismatches:</td> <td>95 </td>	Mismatches:	95
Query Match:	8.07% <td>Indels:</td> <td>128 </td>	Indels:	128
DB:	2	Gaps:	9

US-09-745-506-74 (1-1553) x DB1341 (1-241)

```

OY 245 ATGATTTGAAGGCTTCCTTCTTCTGATGACTTTGATCCCTCTGCTTGTGAG 304
DB 1 MetLysLeuSerGluIleTyrAsnPheLeuAspGlnLeuSerProPheAsnIleGlnGlu 20
OY 305 AGTTGGACAATGTTGATTAAGTGTGAGAACCAAGCCACACATACTGTAATACATC 364
DB 21 SerTyrAspAsnSerGlyIleLeuLeuGlnLysPheArgAspSerGlu-----IleSerThrVal 39
OY 365 TTCCTGACCAATGACCTGACAGAGAGTGTGAGAGAGGTGCTGCAAAAGAGCAGAC 424
DB 40 TyrLeuSerLeuAspIleAspGlnAsnIleLysGlnAlaSerGluAsn-----Ser 57
OY 425 CTCATTTCTTCCTCAATCCGCTATCTTCTGACCACTGAGACCATGACCTGAGACGA 484
DB 58 LeuIleIleThrHisAspProLeuIlePheLysGlyLeuLysAspLeuTyrAspLysThr 77
OY 485 TGAAGAGAGCGCTGTGATCCGGCTCTGAGAAACAGAGTGGATATCTACTCTCCAT 544
  
```



```

Db      78  tyrProalgaIaphelleLysGIuMeIleTyrlYsAnIleSeSerleuIleSerMetHis 97
OY      545  ACAGCCTATGATGCTGCCCCAGGGCGCTCAACACACTGGTGGTAAAGGGCTTGAGCT 604
Db      98  ThrasnTyrrAspLeuSer----- 103
OY      605  TGTACTCTCAGGCCATACATCCTTCACAAAGCTCCCACTACCTACAGAGGAAACGAC 664
Db      103  ----- 103
OY      665  CGAGTGAATTCACAGTTCATACACCAAGACCTGACAAAGTATGTCTGCAGTGA 724
Db      104  -----HisLeuAsnThrTyrrPheThrIuGlu----- 112
OY      725  GGAATTGACGGTGTCTGTCACTTCTTCTGTAGACTGTATAGAGGAACAACA 784
Db      113  -----IleuGluIlyPheLys----- 117
OY      785  CGGATTAACTGGAATTGTACTCAGAGGCTTGATGACAGGTGTAGATTCTTTCCCGG 844
Db      117  ----- 117
OY      845  AACAAACAACCTTATCAGACAGCAATTCCTGCTACGTGAGAACCTTTGCTTACAT 904
Db      118  -----IleSerPheLysAspIlyPheLeuIleTyrr 127
OY      905  ACTGGAATGGACGGTATGACACTGGATGATCTGCTCCCGCAACCATGATTGAT 964
Db      128  -----ValGIuAsnSerMetSerPheGluAlaIleuGlyAsp 139
OY      965  CGAATAAAGACACCTAAACATATCTATATTCGCTTAGCCCTTGGGGTGGGGAACC 1024
Db      140  TrpValIlyLysIleuLeuAsnIleuGlnIleLeuArg-----ValSerAspCys 155
OY      1025  TTAGAGTCTCAAGTAAAGTGTGGCCCTGTGTGCTGTCTGGAGAGCGCTTGTGCAG 1084
Db      156  GlyLysLysAspIleLysArgIleAlaIleCysThrGlySerGlyLysIleuLeuSer 175
OY      1085  GGTGTGAGGCTGACCTTTACCTCAGAGTGCATGATGTCACATGATCTTGGATGCT 1144
Db      176  LysValAspAlaAspCysPheLeuSerGlyAspPheLysTyrrHisGlnAlaIleuGlnAla 195
OY      1145  GCTTCCCAAGCAATTAATGATCTCTGTGTGACACAGACACAGTGAAGAGCGCTTT 1201
Db      196  LeuSerAsnGlnIleSerLeuIleAspLeuGlnHisPheGluSerGluArgTyrrPheSer 215
OY      1202  -----CTTTGTGACCTTGACATATGCTGATTCCTCACTTGAGAAATAAGATTAAT 1252
Db      216  GlnCysLeuAlaIlyAspLeuLysAsnLeu----- 225
OY      1253  AATTATCTATCAGAGACTGACAGGAGCCCTTTCAGTGCTA 1294
Db      226  -----ProLeuGlnValIle 230

```

## RESULT 10

B70777  
 Hypothetical protein RV2230C - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C:Accession: B70777

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; M01D:98293587; PMID:9634230

A:Accession: B70777

A:Status: Preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-379 <COL>

A:Cross-references: GB:I270692; GB:AL123456; NID:93261567; PIDN:CAA94653.1; PID:e235171.

```

A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV2230C

Alignment Scores:
Pred. No.: 4,966-10 Length: 379
Score: 216.50 Matches: 95
Percent Similarity: 44.01% Conservative: 63
Best Local Similarity: 26.46% Mismatches: 158
Query Match: 7.76% Indels: 43
Db: 2 Gaps: 15

US-09-745-506-74 (1-1553) x B70777 (1-379)
OY      299  GCTGAGAGTTGGACAAATGTTGATTACTGTGTGAGAACCAAGCCACCAATAGTAAAT 358
Db      21  AlaGlnSerThrAspSerValGly---LeuValCysGlyAspProAspAlaValAsp 39
OY      359  ACACCTTCCTGACCAATGACCTGACTGAGAGAGTGAATGAGAGAGCTGCTGCAAAAGAG 418
Db      40  SerValThrValAlaValAlaAspAlaThrProAlaValAlaAspGlnValProGln----- 57
OY      419  GCAGACCTCATCTCTCTACCACTCCGCTATCTCCGACCCACGAGACGATACCTGG 478
Db      58  AlaGlyLeuLeuLeuValAlaHisIleProLeuLeuLeuAlaGlyValAspThrValAlaAla 77
OY      479  AACACATGGAAGAGACCGCTGTGATCCGGGCTGTGGAACACAGAGTGTATCTACTCT 538
Db      78  AsnThrProLysGlyValLeuValAlaHisArgLeuIleAlaGlyThrGlyArgSerLeuPheThr 97
OY      539  CCTCATACAGCCTATGATGCTGCGCCCGACAGGCGCTCAACACATGCTGGTAAAGGCTT 598
Db      98  AlaHisThrAsnAlaAspSerAlaSerProGlyValSerAspAlaLeuAlaHisAlaVal 117
OY      599  GGAGCTTGACCTCCAGGCGCCATACATCTTCCAAAGCTCCCAAC----- 643
Db      118  GlyLeuThrValAlaAspAlaValAlaAspPro-----ValProGlyAlaAlaAspLeuAsp 135
OY      644  -----TACCTCAGAGAGGAAACCAACCGAGTA-----GAA 673
Db      136  LysThrValAlaIleTyrrValProArgGlnAsnSerGlnAlaValAlaAlaValAlaPheGln 155
OY      674  TTCAACGTTAACTAACCCCAAGACCTGGACAAAGTATGTCGACGTGAAGGATTT--- 730
Db      156  AlaGlyAlaGlnHisIleGlyAspTyrrSerHisCysSerThrPheValAlaIleTyrrGly 175
OY      731  -----GACGGTGTCTGTGCACTTCTTTTCTGTAGGACTGTGATAG 775
Db      176  GlnPheLeuAlaHisAspIlyAlaSerProAlaIleGlySerValGlyThrValGlnArg 195
OY      776  ---GAACAACAAGGATTAATCTGAATGTGTAAGTGTACTCAGAAAGCTTTGATCAGAGTGTAGAT 832
Db      196  ValAlaGlnAspArgValGlnValAlaAlaProAlaAlaArgAlaGlnValAlaLeuAla 215
OY      833  TTTCTTCCCGGCAACAACAATTAATTCAGAGAG-----GAATTCGTCACTGGAG 886
Db      216  AlaMet---ArgAlaAlaHisProTyrrGluGluProAlaPheAspIlePheAlaLeuVal 234
OY      887  AAGCCTTTCCTCTATACATGTAAGTGGACGGTATGACACACTGATGATCTCTCC 946
Db      235  ProPro-----ProValGlySerGlyLeuGlyAlaGlyIleGlyArgLeuProLysProGluPro 253
OY      947  CTGGACACCATGATGATGCAATTAATAAAGACACCTAAACATGATTCATATTCGTTAGCC 1006
Db      254  LeuArgThrPheValAlaArgLeuGlnAlaAlaAlaProProThr-----Ala 269
OY      1007  CTGGGGGTGGGAGAAACCTTAAGTCTCAAGTCAAGTC-----GTGGCCCTGTGCTGT 1060
Db      270  ThrGlyValAlaArgAlaAlaGlyAspProAspLeuLeuValSerArgValAlaValCysGly 289
OY      1061  GCTTCGGAGGACAGCGCTTCCAGGCTGTGGAGCT-----GACCTTAACCTCAACA 1111
Db      290  GlyAlaGlyAspSerLeuLeuAlaThrValAlaAlaAlaAspValGlnAlaIleTyrrValThr 309

```







Db 133 ----- |||||  
136 GlyIleGlyArg 136  
QY 920 TTTATGCACTGATGATGCTGTCTCCGTCGACCATGATGATGCAATTAAGACAC 979  
Db 137 TleGlyAsnIle---GlnProGlnThrPheTrpGluLeuAlaGlnGlnValLysGlnVal 135  
QY 980 CTAAACATATCTCATTTGCTTACGCTTGGGGTGGGAGAACCTTAGAGTCTCAAGTC 1039  
Db 156 PheAspLeuAspSerLeuArgMetValHisTyrGlnGlnAsnAspGluLysProIle 175  
QY 1040 AAGAGTGGCCCTGTGTGTCTGGAGCAGCCTTCTGCAAGGCTGT----- 1090  
Db 176 SerArgValAlaIleCysGlySerGlyGlnSerPheTyrLysAspAlaLeuAlaLys 195  
QY 1091 GAGCGACCTTACCTACAGCTGAGATGCCATCATCATCTTGTGATGCTGCTCC 1150  
Db 196 GlyAlaAspAlaIleThrGlyAspIleTyrTyrHisThrAlaGlnAspMetLeuSer 215  
QY 1151 CAAAGATTAATGTCATCTCTGTGAACACGACACCTGACAGAGCTTTCTTCTGAC 1210  
Db 216 AspGlyLeuLeuAlaLeuAspProGlyHis---TyrIleGluValLeuPheValGluLys 234  
QY 1211 CTTCGAGATPAGCTGATTTCTCACTTGGAGATTAAG-----ATAAATATTATCTTA 1261  
Db 235 IleAlaAlaLeuLeuThrGlnTrpLysGluLysGlyTyrGluLeuGluIleLeuPro 254  
QY 1262 TCAGAGACTGACAGGACCTCTTCAG 1288  
Db 255 SerGlnAlaSerThrAsnPropheArg 263  
RESULT 14  
T44719  
hypothetical protein MLCB1243.36 [imported] - Mycobacterium lepreae  
C:Species: Mycobacterium lepreae  
C>Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jan-2000  
C:Accession: T44719  
R:Parkhill, J.; Barrall, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, May 1998  
A:Reference number: 222830  
A:Accession: T44719  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-385 <PAR>  
A:Cross-references: EMBL:AL023635; PIDN:CAA19217.1  
C:Genetics:  
A:Experimental source: cosmid B1243  
A>Note: MLCB1243.36  
Alignment Scores:  
Pred. No.: 1 39e-08 Length: 385  
Score: 199.50 Matches: 91  
Percent Similarity: 42.02% Conservative: 59  
Best Local Similarity: 25.49% Mismatches: 170  
Query Match: 7.15% Indels: 37  
DB: 2 Gaps: 15  
US-09-745-506-74 (1-1553) x T44719 (1-385)  
QY 296 TTTGCTGAGAGTTGGACATGTTGATTACTGTGGACCAAGCCACCACTACTGTA 355  
Db 20 PheAlaGlnSerTrpAspSerValGly---LeuValCysGlyAspProGluAspValLeu 38  
QY 356 AATACACTCTCTCTGACCATGACCTGACAGAGAGTGTGAGAGAGGCTGCAAAAG 415  
Db 39 GluAlaIleThrIleLeuValAspAlaThrProAlaValIleAspGluVal-----Pro 56  
QY 416 AAGCAGACCTCATTTCTCTCACTACCTCCCTATCTTCGACCACTGAAGCCGATTAAC 475  
Db 57 AspSerGlyLeuLeuLeuValHisHisProLeuLeuLeuHisGlyValAspThrValAla 76  
QY 476 TGGACACATGAGAGAGGCCCTGGTGATCCGGGCTCTGGAGAACAGAGTCCGTATCTAC 535  
::: ||| ||||| ||| ::::

Db 77 ValSerThrProIysGlyAlaLeuValHisArgLeuIleArgSerGlyArgSerLeuPhe 96  
QY 536 TCTCTCATACAGCCTTATGATGCTGCGCCCGAGGGCTCAACACATGCTGGCTAAAGGG 595  
Db 97 ThrAlaHisThrAsnAlaAspSerLeuSerProGlyValSerAspAlaLeuAlaHisVal 116  
QY 596 CTGGACCTTGTACCTCCAGCCCATCATCT-----TCCAA 634  
Db 117 PheGlyLeuThrValAspAlaValLeuGluProLeuLeuGlyValAlaSerLeuAspLys 136  
QY 635 GCTCCCACTAC---CCTACAGAGGGAACACCGAGTA-----GAATTC 676  
Db 137 TrpValIleTyrValProLeuGlnHisValAlaValAlaGlnAlaValAlaPheGluAla 156  
QY 677 AAGCTTAATCAACCCCAAGACCTGACAAAGTCACTGCTGCGAGGAAGAAAT----- 730  
Db 157 GlyAlaGlnHisIleGlyAspTyrSerHisCysSerTrpSerValIleTrpGlyThrGlyGln 176  
QY 731 -----GAGCGTGTCTCT---GTCACCTCTTTTCTGCTAGACTGTAAATGAG 775  
Db 177 PheMetProHisAspGlyAlaSerProValValGlySerIleGlyAlaIleGluArgVal 196  
QY 776 GAACAACAGGATTAATCTGAATTTACTCAAGAGCTTTGATGACAGGTGGTGAATTT 835  
Db 197 AlaGluAspArgValGluValAlaProAlaArgAlaArgAlaValAlaLeuSerAla 216  
QY 836 CTTTCCCGGAACAACAACCTTATCAAGAAGC---GAAATTCGTACACGGAACGCT 892  
Db 217 MethAlaAlaHisProTyrGluGluProAlaPheAspIlePheAlaLeuValProPro 236  
QY 893 TTGCTTACATACCTGAAATGGAGGTTATGACACACTGATGATGCTGCTCCCTGCA 952  
Db 237 ---ProGlyAspValGlyLeuGlyArgIleGlyThrLeuProArgProGlnSerLeuSer 255  
QY 953 ACCATGATTAATCGATTAATAAAGACACCTAAACTATCTCATTTGCTTAAAGCTTGGG 1012  
Db 256 AlaPheValAlaArgValGlyAlaAlaLeuPro-----GlnThrSerGly 271  
QY 1013 GTGGGAGAACCTTAAGAGTCTCAAGTCAAGTC-----GTGGCCGTGTGCGTGTCT 1066  
Db 272 ValArgAlaThrGlyAspProAspMetLeuValSerArgValAlaValCysGlyGlyAla 291  
QY 1067 GGGAGCAGCTTCTGCAG---GGTGTGAGGCTGACCTT-----TACCTCAGAGTGAG 1117  
Db 292 GlyAspSerLeuLeuSerLeuAlaAlaValAlaAspValGlnAlaTyrValThrAlaAsp 311  
QY 1118 ATGTCCCATCATGATTACTTGGATGCTGCTCCCAAGAAATAAATGATCATCTGTGAA 1177  
Db 312 LeuArgHisHisProAla---AspGluHisArgArgAlaSerAsnValAlaLeuIleAsp 330  
QY 1178 -----CACAGCAACACTGAACAGAGCTTTCTTCTGTACCTTCAGATATGCTGATTTCT 1231  
Db 331 ValAlaHisThrAlaSerGluPheProTyrGlyGlnAlaAlaAspValLeuArgSer 350  
QY 1232 CACTTGAGATAAGATAATTAATTCCTATCAGACACTGACAGGAGACCTT 1282  
Db 351 HisPheGlyThrAlaLeuSerValArgValCysThrIleArgThrAspPro 367  
RESULT 15  
G82901  
conserved hypothetical U0351 [imported] - Ureaplasma urealyticum  
C:Species: Ureaplasma urealyticum  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: G82901  
R:Glass, J.I.; Letkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a  
A:Accession: G82901  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-262 <GLA>  
A:Cross-references: GB:AE002132; GB:AF222894; NID:G6899327; PIDN:AAF30760.1; GSPDB:GN

A: Experimental source: serovar 3; biovar 1  
C: Gene: U0351  
A: Genetic code: SGC3

Alignment Scores:  
Pred. No.: 1.62e-08 Length: 262  
Score: 198.00 Matches: 77  
Percent Similarity: 38.64% Conservative: 59  
Best Local Similarity: 21.88% Mismatches: 104  
Query Match: 7.10% Indels: 112  
DB: 2 Gaps: 12

US-09-745-506-74 (1-1553) x G82901 (1-262)

```

QY 245 ATGATTTGAAGCT-----CTCCCTTCTCTGTAATGACTTTGCATCCCTCTCTCTT 298
DB 4 MetaspilielysAlaGlnAspIleLeuAsnPhelThrLysTyrAspLeuSerLys 23
QY 299 GCTAGAGTTGGACAAATGTTGATTTACTGTGTGAGAACCCACACATATCTGTAAT 358
DB 24 AlaGlnAlaTyrAspLysAsnGlyLeuPhePheAspGln-----GlnThrIleAsn 41
QY 359 ACACTCTCTCCGACCAATGACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 418
DB 42 AsnValGlnIleAlaLeuAspIleThrAspAspValValAsnAspAlaIleLeuAsnAsn 61
QY 419 GCAGACCTCATCTCTCTCTACACCCGCTATCTTCCGACCCATGAGCGCATTAACCTGG 478
DB 62 AlaAsnLeuIleIleSerHisIleProLeuPheThrAsnGlnAspSerAsnAspGluVal 81
QY 479 AACACATGGAAGGAG---CGCCTGTGATCCGGGCTGTGAGAACAGAGTGGTATCTAC 535
DB 82 AsnTyrPheValAsnIleAspLeuIleGluLysIleLysAsnLysIleSerLeuIle 101
QY 536 TCTCCTCATACAGCCTATGATGCTGCGCCAGCGGCTCAACAACTGGTTGGCTAAAGG 595
DB 102 HisLeuHisThrAlaPheAspAlaSerAlaAsnGlyMetSerMetGlnMetAlaLysArg 121
QY 596 CTGGAGCTTTGACCTCCAGGCCCATACATCTTCCAAAGCTCCCACTACCTACAGAG 655
DB 122 LeuGlyLeu----- 124
QY 656 GGAACACACCGAGTAGATTCAGCTTAAGTACACCAAGACCTGGACAAAGTCAATGCT 715
DB 125 -----LeuAsnLeuLysGlnAspGluGlnAsnProTyrLeu 136
QY 716 GCAGTGAAGGAATGAC---GGTGTCTCTGTCACCTTTTCTGTAGAGACTGTAAT 772
DB 137 ValValGlyGluLeuLysLeuGlyValSer----- 146
QY 773 GAGAACAAACACGGATTAACTGAATTGTACTCAGAAAGCTTTGATGCAGGTGTAGAT 832
DB 147 -----ValAsp 148
QY 833 TTTCTTCCCGAACAACACTTTATCAGAGACGGAATTCGTCACTGAGAGAGCT 892
DB 149 TyrIleSerArg-----IleIleLysGlnLysPheLeuSer-----Pro 161
QY 893 TTGCTTACATAGTGAATGAGGAGGTTATGCACACGTGATGATGATCTCTCCCTGGCA 952
DB 162 IleIleLysTyrAsnAsnValPheArg----- 170
QY 953 ACCATGATGTATGATTAAGAACACCTAAACATATTCATATTCGCTTACCCCTGGG 1012
DB 170 ----- 170
QY 1013 GTGGGAGAACCTTGAAGTCTCAAGTCAAGTGTGGCCCTGTGCTGTTCTGGG--- 1069
DB 171 -----LeuGlnThrAsnLeuLysIleGlyIleIleGlyLysSerGlyTyr 186
QY 1070 -----AGCAGGTTCTGAGGGTGTGAGGCTGACCTTACCTCACAGGTGAGAG 1120
DB ----- 1120

```

```

DB 187 LysPheAlaAspAspAlaPheAsnArgTyrGlnLeuAspMetLeuIleThrSerAspLeu 206
QY 1121 TCCCATCATGATGATTTGGATGCTGCTCCCAAGGAATTAATGTCATCTCTGGAACAC 1180
DB 207 LysTyrHisAsnTyrPheAspAlaGlnAlaLysGlnAsnIleIleAspMetAsnHis 226
QY 1181 AGCAACACTGAACAGGCTTCTTCTGACCTTGAGATATGCTG----- 1225
DB 227 Leu---SerGlnSerIlePheIleAspValIleTyrAspGluLeuThrLysPheTyrGly 245
QY 1226 ---GATTCCTCACTTGAGAGATTAATTAATTAATC 1258
DB 246 AsnAspAlaAsnLeuAsnLysSerLeuSerIleIle 257

```

Search completed: August 22, 2003, 14:09:08  
Job time : 75.5 secs